22883

Please type a plus sign (+) inside this box ⊞ PATENT TRADEMARK OFFICE

PTO/SB/69 (12-97)
Approved for use through 9/30/00. OMB 0651-0031
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PETITION ROUTING SLIP

(Find the petition from the list and check the box in the heading above the petition)

Complete if Known	
Application Number	09/093,533
Filing Date	June 8, 1998
First Named Inventor	Malcolm et al
Group Art Unit	2756
Examiner Name	Not Assigned
Attorney Docket Number	CASH-001

☐ PETITIONS DECIDED BY PETITIONS OFFICE

301 Relating to Public Usa Proceedings (37 CFR 1.292)
302 To make application special - prospective manulacture (37 CFR 1.102, MPEP 708.02)
303 To make special - infringement (37 CFR 1.102, MPEP 108.02)
304 Relating to the Filing/Issuance of Divisional Reissue (37 CFR 1.177)
305 To waive or suspend rules (37 CFR 1.183)
306 To expunge a paper from patent application or patent file (37 CFR 1.59)
307 Withdrawal of Attorney (37 CFR 1.36)
308 For access to application except re proceedings before Board (37 CFR 1.14, MPEP 103,104)
309 Refating to Small Entity (37 CFR 1.28)
310 Relating to reexamination (37 CFR 1.181-1.183)
311 For correction of Inventorship lor applications - no filing date (37 CFR 1.48)
312 For correction of inventorship re PCT applications (37 CFR 1.48)
313 For filing application without one or more inventors (37 CFR 1.47)
314 For filing PCT application without one or more inventors (37 CFR 1.47)
315 For extension of time without fee in cases in Application Division (37 CFR 1.136(b))
399 For matters before A/C for Patents - not specified
408 Relating to a filing date under 35 USC 111& 37 CFR 1.53
411 Filing date for application filed by Express Meil (37 CFR 1.10)
412 Filing date for lost application
501 To revive an abandoned application - unavoidable delay (37 CFR 1.137(e))
502 To reviva an abandoned application - unintentional abandonment (37 CFR 1.137(b))
503 To waive/suspend rules (37 CFR 1.183)
504 To invoke supervisory authority - re patent exemining operations (37 CFR 1.181)
505 To withdraw from issue after payment of issue lee (37 CFR 1.313(b)(1-4))
506 To withdraw from issue after payment of issue fee (37 CFR 1.313(b)) or abandon application in lavor ol continuing application
507 To enter priority papers after Issue Fee payment (37 CFR 1.55(a))
508 To defer issuance of patent (37 CFR 1.314)
515 To invoke supervisory authority - re Offica of Admin. (37 CFR 1.181)
516 To waive/suspend rules re patent mailers in Office of Admin. (37 CFR 1.183)
519 To decide matters before Deputy A/C for Patents under 37 CFR 1.182
521 To review refusal to accept & record maintenance lee - application filed on or after 8/27/82 (37 CFR 1.377)
523 To issue patent in the name of the Assignee (37 CFR 1.334(c))
525 To withdraw a holding of abandonment (37 CFR 1.181)
526 To order a Commissioner-initiated Reexamination proceeding (37 CFR 1.520)
527 To convert Provisional Application
528 To reinstate abandoned Provisional Application
530 PCT petition-unavoidable
531 PCT petition-unintentional
532 To accept unavoidably delayed payment of maintenance fee (37 CFR 1.378 (b))
533 To accepl unintentionally delayed payment of maintenance fee (37 CFR 1.378 (c))
534 Petitions related to reexamination proceedings
599 For matters before the Deputy A/C for Patents - not specified

☒ PETITIONS DECIDED BY THE GROUP DIRECTOR

601 To make application speciaf on ground of age or health (37 CFR 1.102, MPEP 708.02)
602 To make special - continuity of earlier application (37 CFR 1.102. MPEP 708.02)
603 To make special - environment quality program (37 CFR 1.102, MPEP 708.02)
604 To make special - accelerated examination (37 CFR 1.102, MPEP 708.02)
605 To make special - Energy Program (37 CFR 1.102. MPEP 708.02)
606 To make special - Recombinant DNA (37 CFR 1.102. MPEP 708.02)
607 To make special lor reasons not provided for in codes 601-606 (37 CFR 1.102, MPEP 708.02)
608 To reopen prosecution after Board decision (37 CFR 1.198)
609 For review of final restriction requirement (37 CFR 1.144)
610 Invoking Authority of Commissioner under 37 CFR 1.181 not specified in codes
611 Relating to the prematureness of final rejection (37 CFR 1.181, MPEP 706,07(c))
612 Relating to the refusal to enter an amendment (37 CFR 1.181 & 1.127, MPEP 714.19)
613 To withdraw a holding of abandonment (37 CFR 1.137, MPEP 711.03)
614 Relating to a requirement to cancel new matter lrom application (37 CFR 1.181. MPEP 608.04(c))
615 Relating to formal sufficiency/propriety of affidavits (37 CFR 1.131, 1.132, 1.608,MPEP 715.07)
616 To institute an interference (37 CFR 1.606)
617 Relating to refusal to enter an amendment under 37 CFR 1.312
618 For concurrent Ex parte and Inter parte proceedings (37 CFR 1.212)
619 For return of original oath of patent application (MPEP 604.04(a))
620 For extension of time (37 CFR 1.136(b))
621 For interviaw after Notice of Allowanca mailed (MPEP 713.10)
622 Concerning appeel application before transler of jurisdiction to Board (MPEP 1206)
623 For second or subsequent suspension of action (37 CFR 1.103, MPEP 709)
625 To reinstate Appeals dismissed in Group
626 From denial of reexamination request (37 CFR 515(c))
627 To enter an amendment after payment of Issue Fee (37 CFR 1.312(b))
628 From refusal to issue a Certilicate of Correction (37 CFR 1.181, MPEP 1480 -1485)
629 For withdrawal of attorney from application pending in group (37 CFR 1.36)
630 For extension of lime in a reexamination (37 CFR 1.550(c))
631 To merge multiple reexamination proceedings (37 CFR 1.565(c), MPEP 2283)
632 To effect a second conversion ol inventorship (37 CFR 1.48, MPEP 201.03)
633 Superconductivity
634 To correct inventorship in e patent not in interference (37 CFR 1.324, MPEP 1481)
635 To change inventorship in an application (37 CFR 1.48)
636 To change inventorship in a patent (37 CFR 1.48)
637 To withdraw from issue before payment of an issue fee (37 CFR 1.313(a))
699 For matters before Group Director - not specified

☐ PETITIONS DECIDED BY BOARD OF PATENT APPEALS AND INTERFERENCES

701 To exercise supervisory authority re action by examiner/examiner-in-chief (37 CFR 1.644)
702 To accept belatedly filed copies of interference sattlement agreements (35 USC 135(c), 37 CFR 1.666(c))
703 For withdrawal of attomey in proceeding under 37 CFR 1.201 - 1.288 (37 CFR 1.36)
704 For access to a settlement agreament under 35 USC 135(c) (37 CFR 1.666(b))
705 For access to an application in proceedings belore the Board (37 CFR 1.14(e))
706 From a refusal to Issue a Certificata ol Correction (37 CFR 1.322, 1.323)
707 To correct errors in inventorship (37 CFR 1.324)
708 For extension of time to file amendment under 37 CFR 1.196(b) (37 CFR 1.136)
709 To make an application before the Board special (37 CFR 1.102)
710 For extension of time to file supplemental Reply Brief (37 CFR 1.136)
711 To assign particular members to hearing or to request augmented panel (35 USC 7)
712 To decide miscellaneous questions in proceedings under 37 CFR 1.601-1.688
713 To accept priority papers in applications in Interference (37 CFR 1.644)
714 To reinstete an Appeal
799 For matters before Chairman of Board - not specified
801 To make an application before the Boerd special (37 CFR 1.102)
802 To reinstata an Appeal
803 To extend time/suspend proceedings (37 CFR 1,196, 1.197, 1.304)
804 For extension of time to file supplemental Reply Brief (37 CFR 1.136)
805 To accapt late request for an Oral Hearing (37 CFR 1.136)
899 For matters befora the Clerk of the Board - not specified

☐ PETITIONS DECIDED BY SPECIAL LAWS (SECURITY AND GOVERNMENT INTEREST MATTERS)

901 Under 42 USC 2182
902 Under 42 USC 2457
903 Under 35 USC 184
904 Under 35 USC 267
905 To consider/review security or Govamment interest matters - not specifled

☐ PETITIONS DECIDED BY THE SOLICITOR

951 Petitions for extension of time in court matters 35 USC 142, 145, 146
952 Petitions reletng to ex parte questions in cases before the Court of Appeals for the Federal Circuit
953 Requests filed under the Freedom of Information Act
959 Not specified

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

CASH-001

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	Art Unit: 2756
Malcolm et al.	Examiner: Not Assigned
Serial No.: 09/093,533	
Filed: June 8, 1998	
For: Network Object Cache Engine	

#13

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in an envelope addressed to:

Assistant Commissioner for Patents
Washington, D.C. 20231

on 1.31.00 (Date). Signed Arlette Malhas
By: Arlette Malhas

Declaration in Support of Accompanying Petition to Make Special

Reason VIII – Special Procedure: Search Was Made

In support of the accompanying Petition to Make Special, applicant's attorney declares as follows:

1. I am the applicant's attorney in the above-identified patent application.

2. The invention of the above application is patentable over the prior art found during an extensive search of the following classes and subclasses of invention for U.S. patents issued up to and including 1/18/2000: 395/200, 395/200.1-200.6, 707/1, 707/10, 709/200, 709/202, 709/203, 709/213, 709/217, 709/218, 709/219, 709/238, 711/112, 711/113, 711/117, 711/124, 711/135, 711/147, and 711/165.

Furthermore, a keyword search was conducted up to January 21, 2000 for all European Patent Office (EPO) issued patents and all EPO pending patent applications of record, and the invention of the above application is patentable over the prior art of record at the EPO. A

keyword search was also conducted up to January 21, 2000 for Japanese Patent Office (JPO) issued patents, and the invention of the above application is patentable over the prior art of record at the JPO.

3. Specifically, the invention provides a method and system for caching information objects transmitted using a computer network. In the invention, a cache engine determines directly when and where to store those objects in a memory (such as random access memory) and mass storage (such as one or more disk drives), so as to optimally write those objects to mass storage and later read them from mass storage, without having to maintain them persistently. The cache engine actively allocates those objects to memory or to mass storage, determines where in mass storage to store those objects, retrieves those objects in response to their network identifiers (such as their uniform resource locators), and determines which objects to remove from the cache so as to maintain appropriate free space.

4. The relevant prior art is summarized and compared to the claimed invention in the following remarks section.

5. I further declare that all statements made herein of my own knowledge are true and that all statements made upon information and belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application and any patent issuing therefrom.

Applicant's undersigned attorney can be reached at (408) 248-1958. All correspondence should continue to be directed to the address indicated below.

Respectfully submitted,

Kevin Roe

Dated: January 31, 2000

Kevin Roe
Reg. No. 40,148

The Swernofsky Law Group
P. O. Box 390013
Mountain View, CA 94039-0013
(650) 947-0700

REMARKS

Detailed Discussion of the Pre-examination Prior Art References

U. S. Pat. No. 5,964,830 to *Durrett* discloses a user portal device for the World Wide Web to communicate with a website server. The user portal device comprises a read only memory (ROM) and a random access memory (RAM), has no physical disk memory, and communicates with a virtual disk server to access software object elements in a non-volatile memory such as a disk drive. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Nor is there any disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile memory. Furthermore, there is no disclosure of optimizing a measure of correlation between the spatial locality of storage of network objects and the temporal locality of retrieval of network objects within a non-volatile disk memory.

U. S. Pat. No. 5,961,602 to *Thompson et al.* discloses a method of retrieving Web content from a plurality of Web servers that the user defines. A test is made to determine whether a given download period has terminated. If the download period has not terminated, a determination is made of an activity level of the communication link. If the activity level is less than a given threshold level, additional requests are issued to the cache according to a "fairness policy" that ensures that content is downloaded from as many servers as possible. However, the present application's claims are not directed to time-shifting the downloading of Web content from multiple servers over a communication link up to a given threshold level. And the present application's claims are not directed to downloading Web content according to a "fairness policy." Furthermore, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects.

U. S. Pat. No. 5,954,795 to *Tomita et al.* discloses a method and apparatus for providing a terminal located on a network, for transferring information desired by a user from an information storage system via a communication line. The terminal comprises storage means for storing information items and information concerning the information items, and information-item determination means for determining, based on the related information, one of the information

items to be provided in response to the user's request. The user selects the information item, based on the related information, according to the information-item determination means. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. There is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile memory. Furthermore, there is no disclosure of optimizing a measure of correlation between the spatial locality of storage of network objects and the temporal locality of retrieval of network objects within a non-volatile disk memory.

U. S. Pat. No. 5,950,205 to *Aviani, Jr.* discloses data transmission over the Internet using a cache memory file system that determines if a requested object is located in cache memory, instead of downloading the object from the originating server. First multimedia objects are written into the cache memory sequentially from the beginning of the cache memory in the order in which they are received. When a first memory amount from a most recently stored one of the first multimedia objects to the end of the cache memory is insufficient to accommodate a new multimedia object, the new multimedia object is written from the beginning of the cache memory, thereby writing over a previously stored one of the first multimedia objects. Later multimedia objects are then written into the cache memory sequentially following the new multimedia object in the order in which they are received, thereby writing over the first ones of the multimedia objects. This cycle is repeated, thereby maintaining a substantially full cache memory. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Furthermore, there is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile memory and selectively deleting old cache entries to make space for new cache entries.

U. S. Pat. No. 5,948,062 to *Tzelnic et al.* discloses a network file server using a cached disk array storing a network file directory. The network file directory includes file locking information and data mover computers, each having file system software for shared read-write file access. The data mover computers perform file system tasks, such as managing the file directory for mapping file names to logical blocks and for locking and unlocking the files. A video file server includes an integrated cached disk array storage subsystem and a plurality of

stream server computers linking the cached disk array to a data network for the transfer of video data streams. The cached disk storage subsystem is responsive to video pre-fetch commands, and the data specified for a pre-fetch command for a process are retained in an allocated portion of the cache memory from the time that the cached disk storage subsystem has responded to the pre-fetch command to the time that the cached disk storage subsystem responds to a fetch command specifying the data for the process. The time between pre-fetching and fetching is selected based on available disk and cache resources. However, pre-fetching data is different from caching data, and this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects.

U. S. Pat. No. 5,946,682 to *Wolfe* discloses a document retrieval system that improves a database system's response time so that a user's request to view new information is serviced quickly. During the time the user spends viewing the displayed information, other information that the user is likely to read or study later is preloaded into memory. If the user does later request this information, this information can be written to the display very quickly because the information need not be retrieved from the database. The present invention takes advantage of the fact that it is possible to accurately predict the information that the user will eventually request by using adaptive prediction schemes. However, preloading information is the same as pre-fetching information, which is different from caching information. Therefore, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Furthermore, there is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile memory.

U. S. Pat. No. 5,905,999 to *Liu et al.*, submitted in a previous IDS, discloses a cache sub-array arbitration circuit for receiving a plurality of address operands from a pending line of processor instructions in order to pre-fetch data needed in any memory access request in the pending instructions. However, this reference is directed to a cache circuit inside a CPU for pre-fetch of data within an individual computer. This reference is not directed to caching a logical group of Internet objects.

U. S. Pat. No. 5,913,033 to *Grout* discloses an apparatus and method for a document manager in a client computer to retrieve documents with embedded links to objects from cache

memory. When a linked object is needed, the document manager parses the embedded link to the object and extracts the file name of the object. The document manager then searches a name list, which contains all the names of objects. If the file name from the link is in the name list, the document manager substitutes a standard-set object of the same name in the local copy of the standard-set. The linked objects that are part of the standard set do not need to be retrieved from the server by the client. However, unlike the present application's claims, this reference requires parsing of the HTML in the documents, extraction of the file name of each object, and substitution of a standard-set object for the embedded linked object. Furthermore, there is no disclosure of optimizing a measure of correlation between the spatial locality of storage of network objects and the temporal locality of retrieval of network objects within a non-volatile disk memory.

U. S. Pat. No. 5,896,506 to *Ali et al.* discloses a method and system for storing and managing objects, such as binary large objects (blobs) in a digital library system. The digital library system includes a plurality of clients, an object server for storing an object, a cache server for storing a copy of the object, and a centralized server for storing information identifying the object as being stored in the object server, associating one or more of the clients with the cache server. One of the clients, as a requesting client, requests retrieval of an object. A copy of the requested object is sent from the cache server to the requesting client if the object is stored in the cache server, and a copy of the object is sent from the object server to the requesting client if the object is not stored in the cache server. A copy of the requested object is sent from the object server to the cache server after the object server sends the object to the client, in which the object sent to the client is made available to the client regardless of whether sending of the copy of the object to the cache server is completed. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Nor is there any disclosure of optimizing a measure of correlation between the spatial locality of storage of network objects and the temporal locality of retrieval of network objects within a non-volatile disk memory. Furthermore, there is no disclosure of caching Internet objects.

U. S. Pat. No. 5,892,937 to *Caccavale*, submitted in a previous IDS, discloses a method and system for dynamically improving the performance of a server in a network. A tuning system monitors the workload of the server in real time, such as the frequency and type of

service requests received by the server from clients in the network. A set of server parameters, such as a data cache hit ratio of a data cache in the server. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects.

U. S. Pat. No. 5,887,151 to *Raz et al.*, submitted in a previous IDS, discloses a method and apparatus of pre-fetching data within a data storage system that includes at least two levels of data storage. The method includes the steps of receiving a pre-fetch command that identifies a list of data blocks within the first level of storage and pre-fetches data from the first level of storage to load data into the second level of storage. However, pre-fetching data blocks is different from caching data blocks, and this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Furthermore, there is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile memory.

U. S. Pat. No. 5,884,046 to *Antonov* discloses a network server that provides the functions of a file server and a local area network switch. A method is also disclosed that includes receiving a message at a first processing node and determining if the message is directed to a workstation computer at a second processing node and sending the message to the second processing node for subsequent delivery to the workstation computer. If the message includes a request for a specific data, a second determination is made if the specific data is in the local cache memory of the first processing node. If the specific data is not in the local cache memory, determination is made if the specific data is in a mass storage device at the first processing node, and if so the mass storage device is accessed to obtain the specific data. If the specific data is not in the local mass storage device, a message is forwarded to a remote processing node to request the specific data. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. There is no disclosure that the steps of maintaining network objects in a cache memory are performed independently of a file system using the cache memory. Furthermore, there is no disclosure of optimizing a measure of correlation between the spatial locality of storage of

network objects and the temporal locality of retrieval of network objects within a non-volatile disk memory.

U. S. Pat. No. 5,878,218 to *Maddalozzo et al.* discloses a method and system for creating and utilizing common caches for inter-networks, using a common cache memory for a private network. A determination is made as to whether a copy of a requested data file is resident in the common cache, and if available, the most recent version is obtained from the common cache. If the requested data file is not available, it is obtained from a server external to the private network. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. There is no disclosure of optimizing a measure of correlation between the spatial locality of storage of network objects and the temporal locality of retrieval of network objects within a non-volatile disk memory.

U. S. Pat. No. 5,870,769 to *Freund*, submitted in a previous IDS, discloses an Internet access system and method that gives a user control over when a document is loaded into a cache. A user is notified of the presence or absence of a document in the cache by changing the visual characteristics of an associated active link status indicator. However, this disclosure is directed to user control of document caching, and this reference does not disclose the present application's claims for using cache memory without user intervention to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects without requiring the intervention of a user. Furthermore, there is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile memory.

U. S. Pat. No. 5,860,106 to *Domen et al.*, submitted in a previous IDS, discloses an apparatus and method for dynamically adjusting the power consumption and performance characteristics of a memory subsystem. By dynamically tracking the behavior of the memory subsystem, the invention predicts the probability the next event will require the attention of a cache memory. However, this reference is directed to dynamic tracking and prediction, and therefore does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects.

U. S. Pat. No. 5,852,717 to *Bhide et al.* discloses systems and methods for performance optimizations for computer networks utilizing HTTP. Performance is increased by reducing the delay period that the client experiences in the time between sending a request to the server and receiving a response. A connection cache is maintained by an agent on the network access equipment to more quickly respond to a request. The agent optionally also maintains a cache of information to more quickly respond to requests for modified objects. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects. Nor is there any disclosure of optimizing a measure of correlation between the spatial locality of storage of network objects and the temporal locality of retrieval of network objects within a non-volatile disk memory.

U. S. Pat. No. 5,819,045 to *Raman et al.*, submitted in a previous IDS, discloses a method for determining a networking capability index, which is representative of the load that a computer applies to a computer network. The method includes selecting packet sizes, measuring the capability of the computer to transmit packets of each of the selected sizes, and determining a probability distribution for the selected packet sizes. However, this reference is directed to computer network load balancing, and therefore does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects.

U. S. Pat. No. 5,802,292 to *Mogul*, submitted in a previous IDS, discloses a method for predictive pre-fetching of objects over a computer network. The method includes predicting in the server computer system a series of subsequent retrieval requests from the client computer system according to predetermined criteria, sending the prediction to the client computer system, and pre-fetching by the client computer system an object based on the prediction. However, pre-fetching objects is different from caching objects. Therefore, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Furthermore, there is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile memory.

U. S. Pat. No. 5,787,470 to *DeSimone et al.* discloses an inter-cache protocol for improving Internet performance that includes querying neighboring caches to determine where a requested object is located for downloading. Requests to the originating server are made when

the requested object is not found or is stale, based on the date and time it has been in the cache. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Nor is there any disclosure of optimizing a measure of correlation between the spatial locality of storage of network objects and the temporal locality of retrieval of network objects within a non-volatile disk memory.

U. S. Pat. No. 5,752,022 to *Chiu et al.* discloses a method for creating a hypertext language for a distributed computer network. A server application intercepts a request for access to documents which may contain embedded references to linked information. The requested document is parsed to discover all embedded addresses, so that the embedded addresses are replaced with the address of the local server. All future references to the embedded information will be supplied by the local server, which implies a persistent cache. Additional linking information are pre-appended and/or post-appended to the requested document by the server. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Furthermore, there is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile disk memory.

U. S. Pat. No. 5,649,154 to *Kumar*, submitted in a previous IDS, discloses a cache memory system with a secondary cache integrated with a direct mapped primary cache in a single structure preferably constructed on a VLSI chip. Data can be read from the secondary cache without loading data first into the direct cache. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Furthermore, there is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile disk memory.

U. S. Pat. No. 5,452,447 to *Nelson et al.* discloses a method and apparatus for a caching file server. The caching file server permits end user client programs on a local node of a computer system to issue requests to read/write data to a remote file and to query/set attributes of the remote file. The cache file server establishes file programs to interface with the client programs, and sets up a common cache for the file attributes. The cache file server also

establishes a memory mapping for the file data which makes use of the Virtual Memory Manager and which establishes its own memory object link to a pager object on the remote node. All file data and attributes are cached locally. The normal Virtual Memory Manager controlled page-in/page outs are still remote and the get attributes and bind operations are performed locally. However, this reference discloses a caching system which extensively communicates with a Virtual Memory Manager. This reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Furthermore, there is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile disk memory.

U. S. Pat. No. 5,452,440 to *Salsburg*, submitted in a previous IDS, discloses a method and structure evaluating and enhancing the performance of cache memory systems, by collecting statistics for quantifying locality of data and thus selecting elements to be cached, and then calculating the overall cache hit rate. This invention provides algorithms that can be directly implemented in software for constructing a precise model that can be used to predict cache hit rates for a cache. However, this reference is directed to optimizing cache memory systems, and therefore does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects.

Discussion of Prior Art Journal Articles

Several of the following journal articles discuss pre-fetching, instead of caching. Internet page retrieval latency reduction by the use of caching techniques or pre-fetching techniques exploit the patterns of Internet page accesses. A caching technique exploits the probability of multiple requests to the same page, but a pre-fetching technique exploits the use of knowledge of a client's requests to multiple pages. Therefore, the prior art references that discuss pre-fetching do not anticipate the present invention's claims concerning caching.

A technical journal article, listed in a previously submitted IDS for this application, entitled "Pre-fetching Links on the WWW," written in 1997 by *Z. Jiang et al.* discusses a pre-fetch scheme including a prediction algorithm and a threshold algorithm. The prediction algorithm collects access history information and estimates the probability with which each file

will be requested in the near future. The threshold algorithm computes the pre-fetch threshold for each server (i.e., a file is pre-fetched if and only if its access probability exceeds its server's pre-fetch threshold). The pre-fetch threshold is dynamically determined based on the system load, capacity and cost of time and resources to the user. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects.

A technical journal article, listed in a previously submitted IDS for this application, entitled "An Interactive Pre-fetching Proxy Server for Improvement of WWW Latency," written in 1997 by *K. Chinen et al.* discusses a pre-fetching scheme called interactive pre-fetching proxy server. The World Wide Web (WWW) is considered to be a page-oriented service, consisting of text, in-line images, and references to other pages. The pre-fetching system gathers references to other pages by parsing the HTML of the client's retrieval page and collects referenced pages at each request of the client. This pre-fetching scheme attempts to reduce the overall latency of page retrieval over the World Wide Web. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects.

A technical journal article, listed in a previously submitted IDS for this application, entitled "The Japan Cache Project: an experiment on domain cache," written in 1997 by *M. Nabeshima* discusses the concept of a domain cache, which is dedicated to handling access to a particular domain name. This caching system attempts to reduce the overall latency of file retrieval over the Internet. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. There is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile memory. Furthermore, there is no disclosure of optimizing a measure of correlation between the spatial locality of storage of network objects and the temporal locality of retrieval of network objects within a non-volatile disk memory.

A technical journal article, listed in a previously submitted IDS for this application, entitled "A Smart Internet Caching System," written in 1996 by *G. Dias et al.* discusses a caching system that attempts to reduce the overall latency of file access over the Internet. A split

gateway with caches on both ends of a slow link and file pre-fetching is proposed for a set of files with a high access probability, called a "neighborhood" of a file. The neighborhood is computed using the previous access patterns for the specific document based on previous users, the known characteristics of HTML, and the general behavior/access patterns of most users. However, there is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile memory. Nor is there any disclosure of optimizing a measure of correlation between the spatial locality of storage of network objects and the temporal locality of retrieval of network objects within a non-volatile disk memory.

A technical journal article, listed in a previously submitted IDS for this application, entitled "Web cache coherence," written in 1996 by *A. Dingle et al.*, discusses the coherence problem of keeping cached pages up to date with the master copies of those pages. The problems of maintaining coherence by using the popular "expiration mechanism" are also discussed. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Furthermore, there is no disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile memory.

A technical journal article, listed in a previously submitted IDS for this application, entitled "Pre-fetching in World Wide Web," written in 1996 by *Z. Wang et al.* discusses the use of pre-fetching in the WWW for reducing the perceived latency. The reduction of latency by the use of caching and the use of pre-fetching both exploit the patterns of Web page accesses. Caching exploits the probability of multiple requests to the same Web page, and pre-fetching exploits the use of the knowledge of a client's requests to multiple Web pages. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects.

A technical journal article, listed in a previously submitted IDS for this application, entitled "A caching relay for the World Wide Web," written in 1994 by *S. Glassman* discusses the design and performance of a caching relay for the World Wide Web. This caching relay handles two protocols (i.e., Gopher and HTTP), and attempts to improve relay performance and reduce its external network traffic. The article mainly discusses experimental measurements of

the relay's performance. However, this reference does not disclose the present application's claims for using cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects. Nor is there any disclosure of caching Internet objects in a non-persistent cache that includes a volatile memory and a non-volatile memory. Furthermore, there is no disclosure of optimizing a measure of correlation between the spatial locality of storage of network objects and the temporal locality of retrieval of network objects within a non-volatile disk memory.

Discussion of How the New Claims Added by Preliminary Amendment Are Distinguishable Over the Pre-examination Search References

In the above-identified application as amended by a preliminary amendment, independent claim 15 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining includes steps of minimizing a measure of latency for displaying a logical group of network objects.

Based on the prior art search previously discussed, claim 15 is not anticipated. The use of cache memory to maintain a logical group of network objects and minimize a measure of latency for displaying the logical group of network objects, is unanticipated by the prior art.

Dependent claims 16-22 are respectfully submitted to be unanticipated for at least the same reasons discussed above with respect to independent claim 15.

Independent claim 23 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining includes steps of optimizing a measure of correlation between (a) spatial locality of storage of said network objects within said mass storage, and (b) temporal locality of retrieval of said network objects.

Based on the prior art search previously discussed, claim 23 is not anticipated. The use of cache memory to optimize a measure of correlation between (a) spatial locality of storage of network objects within mass storage, and (b) temporal locality of retrieval of network objects, is unanticipated by the prior art.

Dependent claims 24-30 are respectfully submitted to be unanticipated for at least the same reasons discussed above with respect to independent claim 23.

Independent claim 31 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining includes steps of determining when and where to record said network objects on said mass storage, in response to a measure of efficiency of said steps of maintaining or serving said network objects.

Based on the prior art search previously discussed, claim 31 is not anticipated. The use of cache memory to determine when and where to record network objects on mass storage, in response to a measure of efficiency of maintaining or serving the network objects, is unanticipated by the prior art.

Dependent claims 32-38 are respectfully submitted to be unanticipated for at least the same reasons discussed above with respect to independent claim 31.

Independent claim 39 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining includes steps of recording said network objects in said memory and retrieving said network objects from said memory, so as to perform at least one of

maximizing a measure of a rate at which said network objects can be written to said mass storage,

maximizing a measure of a rate at which said network objects can be erased from said mass storage,

maximizing a measure of a rate at which said network objects can be retrieved from said mass storage, or

minimizing a measure of latency time for retrieving said network objects from said mass storage.

Based on the prior art search previously discussed, claim 39 is not anticipated. The use of cache memory to include steps of recording network objects and retrieving network objects, so as to maximize a measure of a rate at which network objects can be written to mass storage, is unanticipated by the prior art. The use of cache memory to include steps of recording network objects and retrieving network objects, so as to maximize a measure of a rate at which network objects can be erased from mass storage, is unanticipated by the prior art. The use of cache memory to include steps of recording network objects and retrieving network objects, so as to perform maximizing a measure of a rate at which network objects can be retrieved from mass storage, is unanticipated by the prior art. The use of cache memory to include steps of recording network objects and retrieving network objects, so as to minimize a measure of latency time for retrieving network objects from mass storage, is unanticipated by the prior art.

Dependent claims 40-46 are respectfully submitted to be unanticipated for at least the same reasons discussed above with respect to independent claim 39.

Independent claim 47 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining is performed independently of a file system using said mass storage.

Based on the prior art search previously discussed, claim 47 is not anticipated. The use of cache memory to maintain network objects independently of a file system using mass storage, is unanticipated by the prior art.

Independent claim 48 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining includes steps of selecting a group of more than one said network objects to be written to said mass storage collectively, and writing said group of network objects to said mass storage in one or more write episodes.

Based on the prior art search previously discussed, claim 48 is not anticipated. The use of cache memory for selecting a group of more than one network objects to be written to mass storage collectively, and writing the group of network objects to the mass storage in one or more write episodes, is unanticipated by the prior art.

Independent claim 49 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining includes steps of writing a group of network objects to said mass storage in one or more write episodes, such that at least one parameter of said write episodes is responsive to a measure of efficiency of said steps of maintaining or serving said network objects.

Based on the prior art search previously discussed, claim 49 is not anticipated. The use of cache memory for writing a group of network objects to mass storage in one or more write

episodes, such that at least one parameter of the write episodes is responsive to a measure of efficiency of the steps of maintaining or serving the network objects, is unanticipated by the prior art.

Independent claim 50 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining includes steps of selecting a group of more than one said network objects to be deleted from said mass storage collectively, and deleting said group of network objects to said mass storage in one or more delete episodes.

Based on the prior art search previously discussed, claim 50 is not anticipated. The use of cache memory for selecting a group of more than one said network objects to be deleted from mass storage collectively, and deleting the group of network objects to the mass storage in one or more delete episodes, is unanticipated by the prior art.

Independent claim 51 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining includes steps of deleting a group of network objects from said mass storage in one or more delete episodes, such that at least one parameter of said delete episodes is responsive to a measure of efficiency of said step of maintaining or serving said network objects.

Based on the prior art search previously discussed, claim 51 is not anticipated. The use of cache memory for deleting a group of network objects from the mass storage in one or more delete episodes, such that at least one parameter of the delete episodes is responsive to a measure of efficiency of the step of maintaining or serving the network objects, is unanticipated by the prior art.

Independent claim 52 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said cache memory includes at least a portion thereof that is non-persistent.

Based on the prior art search previously discussed, claim 52 is not anticipated. The use of a cache memory that includes a portion of the cache memory that is non-persistent for a set of network objects, is unanticipated by the prior art.

Independent claim 53 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining includes steps of recording said network objects in said memory and retrieving said network objects from said memory, without having to maintain said network objects persistently.

Based on the prior art search previously discussed, claim 53 is not anticipated by any issued U. S. patent, issued or pending EPO patent of record, or Japanese patent of record. The use of a cache memory for recording and retrieving a set of network objects, without maintaining network objects persistently, is unanticipated by the prior art.

Independent claim 54 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining includes a step of writing a group of network objects to said mass storage in one or more write episodes, such that said write episodes

are performed so as to atomically commit changes to said mass storage during each said write episode, and whereby said information remains available after at least one of (a) loss of power, (b) loss of storage, or (c) immediate failure of at least a portion of said memory.

Based on the prior art search previously discussed, claim 54 is not anticipated. The use of a cache memory for writing a group of network objects to mass storage in one or more write episodes, such that the write episodes are performed so as to atomically commit changes to the mass storage during each write episode, and the information remains available after at least one of (a) loss of power, (b) loss of storage, or (c) immediate failure of at least a portion of the memory, is unanticipated by the prior art.

Independent claim 55 states a method including the elements of:

receiving a set of network objects in response to a first request to an information provider from an information requester; and

maintaining said network objects in a cache memory, said cache memory including mass storage;

wherein said step of maintaining includes a step of deleting a group of network objects to said mass storage in one or more delete episodes, such that said delete episodes are performed so as to atomically commit changes to said mass storage during each said delete episode, and whereby said information remains available after at least one of (a) loss of power, (b) loss of storage, or (c) immediate failure of at least a portion of said memory.

Based on the prior art search previously discussed, claim 55 is not anticipated. The use of a cache memory for deleting a group of network objects from mass storage in one or more delete episodes, such that the delete episodes are performed so as to atomically commit changes to the mass storage during each delete episode, and the information remains available after at least one of (a) loss of power, (b) loss of storage, or (c) immediate failure of at least a portion of the memory, is unanticipated by the prior art.

SUMMARY

A pre-examination search was made of patents of record in the U. S., Europe and Japan. According to the pre-examination search results, the new pending claims 15-55, as specified in a preliminary amendment for the above-identified application, are not anticipated by the prior art. Applicant's attorney respectfully requests advancement of examination of the above-identified application under 37 CFR §1.102 and MPEP §708.02.

Applicant's undersigned attorney can be reached at (408) 248-1958. All correspondence should continue to be directed to the address indicated below.

Respectfully submitted,

Kevin Roe

Dated: January 31, 2000

Kevin Roe
Reg. No. 40,148

The Swernofsky Law Group
P. O. Box 390013
Mountain View, CA 94039-0013
(650) 947-0700